Rešitve kolokvija iz verjetnosti in statistike z dne 27. 1. 2010

IŠRM

1. Ker je Y diskretna slučajna spremenljivka, ki zavzema vrednosti iz $\mathbb{N}$, lahko uporabimo izrek o popolni verjetnosti:

$$
\begin{aligned}
P\left(\frac{X}{Y} > 1\right) &= \sum_{k=1}^{\infty} P(Y = k)\, P\left(\frac{X}{Y} > 1 \,\middle|\, Y = k\right) = \\
&= \sum_{k=1}^{\infty} P(Y = k)\, P\left(\frac{X}{k} > 1\right) = \\
&= \sum_{k=1}^{\infty} P(Y = k)\, P(X > k) = \\
&= \sum_{k=1}^{\infty} \frac{\frac{1}{2} - \Phi(k)}{2^k}
\end{aligned}
$$

Zadostuje sešteti prva dva člena, rezultat je $0'085$.

2. Uporabimo lahko centralni limitni izrek v šibkejši ali močnejši različici.

- V močnejši različici ga lahko uporabimo z argumentom, da je S vsota slučajnih spremenljivk $X_1, \ldots X_{100}$, $2X_{101}, \ldots 2X_{200}$, ki jih je veliko, so neodvisne in nobena ne izstopa posebej.
- V šibkejši različici pa ga lahko uporabimo:
  - bodisi z argumentom, da je S vsota slučajnih spremenljivk $X_1 + 2X_{101}, \ldots,$ $X_{100} + 2X_{200}$, ki jih je veliko, so neodvisne in enako porazdeljene;
  - bodisi z argumentom, da je S vsota slučajnih spremenljivk $X_1 + \ldots + X_{100}$ in $2X_{101} + \ldots + 2X_{200}$, ki sta obe vsoti veliko neodvisnih in enako porazdeljenih slučajnih spremenljivk ter zato približno normalni.

Iz:

$$
\begin{aligned}
E(X_i) &= \int_0^{\infty} x\, e^{-x}\, \mathrm{d}x = 1\,, \\
E(X_i^2) &= \int_0^{\infty} x\, e^{-x}\, \mathrm{d}x = 2\,, \quad D(X_i) = 1
\end{aligned}
$$

izračunamo:

$$
E(S) = (100 + 100 \cdot 2) \cdot 1 = 300\,, \quad D(S) = (100 + 100 \cdot 2^2) \cdot 1 = 500\,.
$$

Sledi:

$$
P(S > 300) \approx \frac{1}{2} - \Phi\left(\frac{330 - 300}{\sqrt{500}}\right) \doteq 0'0899\,.
$$

3. Iz funkcije verjetja:

$$L = 4\ln a + 5\ln b + \ln(1 - 2a - b)$$

in njenih parcialnih odvodov:

$$\frac{\partial L}{\partial a} = \frac{4}{a} - \frac{2}{1 - 2a - b}, \quad \frac{\partial L}{\partial b} = \frac{5}{b} - \frac{1}{1 - 2a - b}$$

dobimo oceno $\hat{a} = 0\text{˙}2$, $\hat{b} = 0\text{˙}5$. Ocenjena porazdelitev je torej:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0\text{˙}2 & 0\text{˙}2 & 0\text{˙}5 & 0\text{˙}1 \end{pmatrix}.$$

4. Označimo z z kvantil standardne normalne porazdelitve za verjetnost $0\text{˙}975$, z $n = 100$ velikost vzorca, s $\hat{p}$ pa standardno točkasto oceno za p, t. j. $\hat{p} = 0\text{˙}6$.

a) Po standardni približni metodi iz:

$$\Delta = \frac{z\sqrt{\hat{p}(1 - \hat{p})}}{\sqrt{n}} \doteq \frac{1\text{˙}96 \cdot \sqrt{0\text{˙}6 \cdot 0\text{˙}4}}{\sqrt{100}} \doteq 0.0961$$

(zaokroženo navzgor) dobimo interval zaupanja $0\text{˙}5039 < \Delta < 0\text{˙}6961$.

b) Hipoteze ne zavrnemo, če velja:

$$\frac{|\hat{p} - p_0|}{\sqrt{p_0(1 - p_0)}}\sqrt{n} < z\,.$$

Za rešitev te neenačbe najprej opazimo, da gre izraz na levi proti neskončno, brž ko gre p_0 proti 0 ali 1. Nato rešimo ustrezno enačbo, ki je ekvivalentna enačbi:

$$\frac{n(\hat{p} - p_0)^2}{p_0(1 - p_0)} = z^2$$

oziroma (za $0 < p_0 < 1$):

$$(n + z^2)p_0^2 - (2n\hat{p} + z^2)p_0 + np_0^2 = 0$$

Rešitvi sta:

$$p_0 = \frac{2n\hat{p} + z^2 \pm \sqrt{4n\hat{p}(1 - \hat{p})z^2 + z^4}}{2(n + z^2)}$$

Hipoteze torej ne zavrnemo pri tistih p_0, ki ležijo med obema rešitvama. Zahtevana vrednost bo torej zgornja rešitev:

$$p_0 = \frac{2n\hat{p} + z^2 + \sqrt{4n\hat{p}(1 - \hat{p})z^2 + z^4}}{2(n + z^2)} \doteq 0\text{˙}6906\,.$$